Dynamic Domain Specificity In Human Ventral Temporal Cortex

Brett B. Bankson[1,2,4*], Matthew J. Boring[1,3,4], R. Mark Richardson[5,6], Avniel Singh Ghuman[1,2,3,4]

[1]Laboratory of Cognitive Neurodynamics, Department of Neurological Surgery, University of Pittsburgh Medical Center, 200 Lothrop Street, Pittsburgh, PA, 15213.

[2]Department of Psychology, University of Pittsburgh, 210 South Bouquet St, Pittsburgh, PA 15260, USA.

[3]Center for Neuroscience at the University of Pittsburgh, University of Pittsburgh, A210 Langley Hall, Pittsburgh, PA 15260, USA.

[4]Center for the Neural Basis of Cognition, 4400 Fifth Avenue, Pittsburgh, PA 15213.

[5]Department of Neurosurgery, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02144.

[6]Harvard Medical School, 25 Shattuck St., Boston, MA 02115.

Abbreviated title: Temporal dynamics of object individuation

*Correspondence to:
Brett B. Bankson
Laboratory of Cognitive Neurodynamics
UPMC Presbyterian
Suite B-400
200 Lothrop Street
Pittsburgh, PA 15213
bbb17@pitt.edu

Number of Pages 38; Figures 5
Number of Words 8934 Abstract 217; Introduction 648; Discussion 1126
The authors declare no competing interests.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health (R01MH107797 and R21EY030297 to A.G) and the National Science Foundation (Graduate Research fellowship to B.B.B., 1734907 to A.G.). We would like to thank the patients, their families, and nurses, staff, and physicians at the Epilepsy Monitoring Unit and the University of Pittsburgh Comprehensive Epilepsy Center at the University of Pittsburgh Medical Center, without whom this study would not be possible. We would also like to thank Michael Ward, Sean Walls, and Ellyanna Kessler for assistance in data collection and Julie Fiez for assistance with design of the word experiment. Additional thanks to Chris Baker, Brad Mahon, and Alex Martin for critical comments and feedback on this work.

ABSTRACT

An enduring neuroscientific debate concerns the extent to which neural representation is restricted to networks of patches specialized for particular domains of perceptual input (Kaniwsher et al., 1997; Livingstone et al., 2019), or distributed outside of these patches to broad areas of cortex as well (Haxby et al., 2001; Op de Beeck, 2008). A critical level for this debate is the localization of the neural representation of the identity of individual images, (Spiridon & Kanwisher, 2002) such as individual-level face or written word recognition. To address this debate, intracranial recordings from 489 electrodes throughout ventral temporal cortex across 17 human subjects were used to assess the spatiotemporal dynamics of individual word and face processing within and outside cortical patches strongly selective for these categories of visual information. Individual faces and words were first represented primarily only in strongly selective patches and then represented in both strongly and weakly selective areas approximately 170 milliseconds later. Strongly and weakly selective areas contributed non-redundant information to the representation of individual images. These results can reconcile previous results endorsing disparate poles of the domain specificity debate by highlighting the temporally segregated contributions of different functionally defined cortical areas to individual level representations. Taken together, this work supports a dynamic model of neural representation characterized by successive domain-specific and distributed processing stages.

SIGNIFICANCE STATEMENT

The visual processing system performs dynamic computations to differentiate visually similar forms, such as identifying individual words and faces. Previous models have localized these computations to 1) circumscribed, specialized portions of the brain, or 2) more distributed

aspects of the brain. The current work combines machine learning analyses with human intracranial recordings to determine the neurodynamics of individual face and word processing in and outside of brain regions selective for these visual categories. The results suggest that individuation involves computations that occur first in primarily highly selective parts of the visual processing system, then later recruits highly and non-highly selective regions. These results mediate between extant models of neural specialization by suggesting a dynamic domain specificity model of visual processing.

INTRODUCTION

A key debate regarding the architecture of the cortex concerns the extent to which diagnostic aspects of stimuli are processed within circumscribed, domain-specific and potentially modular cortical networks (Kaniwsher et al., 1997; Martin, 2007; Livingstone et al., 2019), or distributed across large, overlapping sections of cortex as a feature map (Haxby et al., 2001; Op de Beeck, 2008). On one hand, an extensive body of primate single unit recordings (Perrett et al., 1984; Tsao et al., 2006), human neuroimaging (Kanwisher et al., 1997; Puce et al., 1996), stimulation (Puce et al., 1999; Hirshorn et al., 2016; Parvizi et al., 2012; Afraz et al., 2006; Pitcher et al., 2007), and lesion studies suggests that perception is causally related to the activity within patches of cortex that respond selectively to preferred stimulus categories (Farah et al., 1995; Schalk et al., 2017). Conversely, the distributed feature map hypothesis is supported by evidence from both neuroimaging and single unit recordings that shows reliable face differentiation outside of strongly face-selective patches (Haxby et al, 2001; Bell et al., 2011) and differentiation of non-face categories within face patches (Kiani et al., 2007; Cukur et al., 2013; Hanson & Schmidt, 2011). This hypothesis posits that category selective patches are clusters that

process visual information that is particularly diagnostic for those categories within a larger, continuous feature map that spans ventral temporal cortex (VTC) (Op de Beeck et al., 2008; Mur et al., 2012).

Across these hypotheses, a central point of debate concerns the role of activity evoked by stimuli outside of highly selective patches (e.g. face-related activity outside of face patches) and activity evoked by "other" stimuli inside patches selective for particular categories of stimuli (e.g. non-face activity inside face patches). Specifically, a critical tension between the aforementioned hypotheses is whether individual-level discrimination (e.g. recognizing which face or word a person is viewing) can be found outside of putative category-selective patches (Spiridon & Kanwisher, 2002; Nestor et al., 2011). Examining individual level discrimination is crucial because it probes the potential computational role and representational level of activity inside and outside of patches that are highly selective for stimuli at the category level. Indeed, the sparing of category-level discrimination in various agnosias (Damasio et al., 1982) emphasizes that individual representations are a key level in the debate between domain specific and distributed models of processing.

To test for the presence of individual-level representations across time in and out of selective patches, the dynamics of face individuation was examined with intracranial electroencephalography (IEEG) in 14 patients with pharmacologically intractable epilepsy. To ensure that face individuation was based on face identity level and not the visual image level, 15 different images of each of 14 different identities were used across 5 expressions (anger, sadness, fear, happy, neutral) and 3 gaze directions (left, straight, right). The dynamics of word individuation was examined in 5 additional patients (2 overlapping, 17 total patients in the study). Electrode contacts of interest were restricted anatomically to VTC, inferior to the middle

temporal sulcus. Using category localizer data, these contacts were first functionally partitioned into ones that demonstrated high face selectivity (HFS) or high word selectivity (HWS) and those with low face selective (LFS) and low word selective (LWS) responses (Figure 1). Every patient had recordings from both high and low category-selective areas. Once the contacts had been partitioned according to their category selectivity, machine learning analyses (regularized logistic regression) were used to compare the dynamics of individuation within and outside of highly category-selective areas. Above chance classification of individual faces and words was seen in both high and low face and word selective regions, but significant decoding emerged approximately 170 ms earlier in high selectivity regions compared to low selectivity regions. These results suggest a dynamic model of domain specificity in VTC in which processing is first restricted to highly selective regions and then is processed by both high and low selectivity regions.

MATERIALS AND METHODS

Subjects

Experimental protocols were approved by the Institutional Review Board of the University of Pittsburgh and written informed consent was obtained from all subjects. 17 patients (8 female) undergoing surgical treatment for medicine-resistant epilepsy volunteered to participate in this experiment. Patients had previously undergone surgical placement of subdural electrocorticographic contacts or stereoelectroencephalography (collectively referred to as iEEG here) as standard care for clinical monitoring during seizure onset zone localization. The ages of subjects ranged from 20 to 64 years (mean = 39.1, SD = 14.6). None of the subjects showed any ictal events during experimental recording nor epileptic activity on the electrodes used in this

study. All patients completed a localizer session, 14 patients completed experiment 1, and 5 patients (2 overlap) completed experiment 2.

Experimental Design: Stimuli

In the localizer session, images of bodies (50% male), faces (50% male), words, hammers, houses and phase scrambled faces were used. Examples of these stimuli are outlined in Figure 2 of Ghuman et al. (2014). Phase scrambled images were created in Matlab by taking the two-dimensional spatial Fourier spectrum of the image, extracting the phase, adding random phases, recombining the phase and amplitude, and taking the inverse two-dimensional spatial Fourier spectrum. Each image category was presented 80 times, yielding a total of 480 image presentations. Each image was presented in pseudorandom order and repeated once in each session.

For experiment 1, frontal views of 14 different face identities were drawn from the Radboud Faces Database. 15 images of each identity were presented, with 5 expressions (anger, sadness, fear, happy, neutral) and 3 gaze directions (left, right, forward). Each unique image was presented four times, yielding a total of 60 presentations per identity and 840 face image presentations. For experiment 2, 36 different three and four-letter real, pseudo-, false font words were presented 30 times each, yielding a total of 1080 word presentations. Only data from trials corresponding to four-letter real and pseudo-words were considered further for data analysis, corresponding to 16 unique words. Word stimuli were selected to have similar log frequency, mean bigram frequency and bigram frequency by position across similar and dissimilar word pairs (measured using the English Lexicon Project). All stimuli for the three experimental sessions were presented on an LCD computer screen placed ~1 meter from subjects' heads.

Experimental Design: Paradigms

In all experimental sessions, each image was presented for 900 ms with 900 ms inter-trial interval during which a fixation cross was presented at the center of the screen (~10° × 10° of visual angle for the localizer session and experiment 1, ~6° × 6° visual angle for experiment 2). For the localizer session, images were repeated 20% of the time at random. Subjects were instructed to press a button on a button box when an image was repeated (1-back). Only the first presentations of repeated images were used in the analysis.

In experiment 1, subjects completed a gender discrimination task, reporting whether the presented face was male or female via button press on a button box. Each subject completed one or two sessions of the task. All three paradigms were coded in MATLAB (version 2007, Mathworks, Natick, MA) using Psychtoolbox (Brainard, 1997) and custom written code.

In experiment 2, subjects completed a lexical decision task, reporting whether the presented word was real or not (pseudoword and false font comparisons) via button press on a button box. Each subject completed one or two sessions of the task. All three paradigms were coded in MATLAB using Psychtoolbox and custom written code.

Data preprocessing

Electrophysiological activity was recorded at 1000 Hz using iEEG electrodes. Single-trial potential was extracted by band-pass filtering the raw data between 0.2-115 Hz using a fourth-order Butterworth filter to remove slow drift, high-frequency noise, and 60 Hz line noise (additionally using a 55-65 Hz stop-band). For each trial, the power spectrum density (PSD) at 2-100 Hz with a bin size of 2 Hz and time-step size of 10 ms was estimated using a Hann multi-

177 taper power spectrum analysis in the FieldTrip toolbox (Oostenveld et al., 2011). For each
178 channel, the neural activity between 50-300 ms prior to stimulus onset was used as baseline, and
179 the PSD at each frequency z-scored based on the mean and variance of baseline activity.
180 Broadband gamma signal was extracted as mean z-scored PSD across 40-100 Hz. Both the single
181 trial potentials (stP) and single trial broadband high-frequency activity (stBHA) were used in all
182 analyses.

183 Trials with a maximum amplitude five standard deviations above the mean across trials
184 were eliminated, as well as trials with a deflection greater than 25 μV between sampling points.
185 These criteria allow the rejection of sampling error or ictal events, and resulted in elimination of
186 less than 1% of trials when applied in this and previous work (Li et al., 2019).

187

188 **Electrode localization**

189 To accurately identify electrode contact location, the co-registration of grid electrodes and
190 electrode strips with cortex was adapted from Hermes et al. (2017). Electrode contacts were
191 segmented from high-resolution post-operative computerized tomography (CT) scans of patients
192 and co-registered with anatomical MRI scans that were conducted before neurosurgery and
193 electrode implantation. This method accounted for shifts in specific electrode location caused by
194 potential deformation of the cortex that arise when utilizing FreeSurfer
195 (https://surfer.nmr.mgh.harvard.edu/, 1999) software reconstructions to co-register with the CT
196 scans. SEEG electrodes were localized with Brainstorm software (Tadel et al., 2011) that co-
197 registers post-operative MRI with pre-operative MRI images. Complete localization
198 (incorporating the following electrode selection step) is depicted in Figure 1. The presence of
199 numerically greater HFS contacts in the left hemisphere than right hemisphere is most likely

200 explained by the larger absolute numbers of left than right hemisphere electrode contacts, a result
201 of electrode placement being guided solely by clinical needs of each patient.
202
203 **Electrode selection**
204 Electrodes were selected according to anatomical and two functional criteria. Anatomically,
205 electrodes of interest were selected from within ventral temporal cortex below the middle
206 temporal gyrus. Specifically, the midline of the middle temporal gyrus was defined as the upper
207 limit for anatomical consideration: the beginning of the middle temporal gyrus was used to
208 define a posterior threshold, and the midline of the middle temporal gyrus terminating at the
209 temporal pole was used as the anterior threshold for electrode selection. We conducted
210 multivariate classification over data from the localizer session to identify face and word sensitive
211 electrodes (described in next section). Functionally, highly category selective electrodes of
212 interest demonstrated a peak six-way face classification *d'* score greater than 0.8, corresponding
213 to $p < .01$ and a large effect size (Cohen, 1988) for the preferred category (face or word) using a
214 Naïve Bayes classifier. Electrodes were not considered if a *d'* score greater than 0.8 resulted
215 from systematically lower face or word sTP values relative to other conditions (whereby above
216 chance classification could occur simply by systematically lower response magnitude). Selective
217 electrodes were also required to show a maximal sTP or stBHA response to either faces or words
218 for at least 50 ms during the stimulus presentation period. Within each patient's montage, all
219 VTC electrodes of interest that did not meet the criteria for high selectivity for faces or words
220 were labeled as low selectivity (note that face selective electrodes could be considered low
221 selectivity for words and vice versa). Finally, to control for any systematic differences in
222 anatomic location between high selectivity and low selectivity contacts, the most anterior low

223 selectivity contacts from each montage (which were more numerous and more anteriorly located,
224 in general, than the selective contacts) were removed until the high selectivity and low selectivity
225 contacts from each montage were matched anatomically along the anterior-posterior axis.
226 Functionally, this trimming procedure yielded high selectivity and low selectivity contact
227 populations in each patient's montage with equivalent mean coordinate values along the anterior-
228 posterior axis and ensured that any latency differences between populations could not
229 immediately be attributed to any expected conduction delays. Indeed, recent work has
230 demonstrated a relationship between response onset latency and situation along the anterior-
231 posterior axis, such that more anterior contacts emerged later in time (Schrouff et al., 2020).
232 Note that this anatomical balancing procedure did not meaningfully alter the time course of
233 classification over low selectivity contacts compared to retaining all anterior low selectivity
234 contacts and all results remained similar if non-balanced electrodes were used in the analyses.
235

236 **Multivariate classification: Naïve Bayes classifier**

237 We first used a Naïve Bayes classifier with 3-fold cross validation to examine category
238 selectivity over time at individual electrode contacts throughout ventral temporal cortex. Both
239 sTP and stBHA signal values were used as input features in the classifier with a sliding 100 ms
240 time window (10 ms width) as previous studies have shown increased sensitivity and specificity
241 when using both sTP and stBHA (Miller et al., 2016). Indeed, sTP and stBHA metrics have been
242 shown to capture separate and complementary aspects of the physiology that contribute to visual
243 processing as measured with iEEG (Lescynski et al., 2019). sTP signal was sampled at 1000 Hz
244 and stBHA at 100 Hz, which yielded 110 features (100 mean sTP voltage potentials and 10
245 normalized mean stBHA PSD values). Thus at each time point at each electrode, the classifier

was trained on the first 2-folds and performance evaluated on the left out fold for 6-way classification of the six object categories presented in the localizer session. We used the sensitivity index *d'* for face or word category against all other categories to determine face and word selective contacts. *d'* was calculated as Z(true positive rate) – Z(false positive rate), where Z is the inverse of the Gaussian cumulative distribution function.

Elastic net regularized logistic regression

To examine the temporal dynamics of face and word individuation, we used elastic net regularized logistic regression with three-fold cross validation implemented with the GLMNET package in Matlab. Elastic net was chosen as a means to identify diagnostic electrode contacts by removing non-informative and/or highly correlated classifier features. These series of classification problems were conducted iteratively in four different electrode populations: individual face classification from experiment 1 data in VTC high face selective contacts and VTC low face selective contacts, and individual word classification from experiment 2 data in VTC high word selective contacts and VTC low word selective contacts. Face identity classification was conducted across expression and gaze direction, effectively varying the low-level visual features of each face identity such that this classification problem was not simply face image classification.

sTP signal was first downsampled to 100 Hz to yield an equal amount of sTP and stBHA features. sTP signal was then normalized with a Box-Cox transformation to enhance interpretability of classifier weights. Thus at each time point, sTP and stBHA values from each trial were arranged as a *P*-dimensional vector corresponding to 2 * number of contacts in each of the four predefined electrode contact populations. The time course of face and word

individuation was identified by examining the pairwise decoding accuracy of a classifier using 3-fold cross-validation. The regularization parameter (λ) was set a priori to 0.9 to favor more sparse classification solutions. The results of this analysis are depicted in Figure 2. For display purposes, group mean time courses were smoothed with a moving average of 30 millisecond fixed window length.

For comparison purposes, L1 regularized logistic regression was also repeated in the same manner as the above elastic net analyses (classification conducted separately for high and low category selectivity populations) to demonstrate minimal difference in the time course of *d'* values from the different regularization procedures. For this analysis, the regularization parameter (λ) was by default set to 1.

To demonstrate the robustness of general trends of individuation across high and low selectivity contact populations, the elastic net classification procedure was repeated with additional thresholds determined by dividing face and word contact populations into partitions of equal numbers. To do so, all contacts across all subjects in face and word tasks, respectively, were sorted according to peak d' selectivity value from the category localizer. Then, these contacts were divided into six equal partitions. For example, contacts in the face task were divided into 6 partitions consisting of 71 contacts each. Then, elastic net regularized classification was conducted again according to the following groupings: 1) bottom two partitions labeled as LFS, top four partitions labeled as HFS (corresponding d' value of 0.61 dividing the two groups); 2) bottom three partitions labeled as LFS, top three partitions labeled as HFS (corresponding d' value of 0.7 dividing the two groups); 3) bottom four partitions labeled as LFS, top two partitions labeled as HFS (corresponding d' value of 0.82 dividing the two groups). This procedure was repeated for word selective contacts at the following d' thresholds:

292 0.58, 0.67, 0.86. The results of this analysis are depicted in Figure 3. For display purposes, group
293 mean time courses were smoothed with a moving average of 30 millisecond fixed window
294 length. The partitions corresponding to the bottom 1/6 and top 5/6 (and vice versa) are not
295 demonstrated because not all subjects had contacts in the lowest and highest partitions.
296

297 **L1 regularized logistic regression**

298 To examine the diagnosticity of brain activity from high and low category selectivity electrode
299 populations in concert with one another, we repeated the above classification analyses with L1 as
300 opposed to elastic net regularization and examined the proportion of electrode contacts that were
301 entirely penalized and removed from the classifier model. Additionally, all VTC electrode
302 contacts (high and low category selectivity) were used to train each classifier, as opposed to
303 splitting the electrode populations as in the previous analyses. After conducting pairwise face
304 classification and pairwise word classification, the classifier weights from each pairwise
305 classification for each electrode contact were extracted and the number of non-zero (positive or
306 negative) weights for each contact tabulated. The percent of electrode contacts with non-zero
307 weights was determined at every time point after baseline normalization. Baseline normalization
308 consisted of determining the threshold of non-zero weight counts that would yield <1% contacts
309 with non-zero weights during the baseline period. The total percentage of electrode contacts
310 assigned non-zero weights for at least 50 ms across the entire time course was determined, and
311 results from this analysis are depicted in Figure 5A. This change in classifier does not alter the
312 time course of individuation compared to the original elastic net procedure.
313

314 **Electrode Diagnosticity in low category selectivity areas**

315 Having examined the contributions of high and low face and word selectivity contacts to
316 exemplar representation, we were then interested in examining whether low face and word
317 selective sites that demonstrated selectivity for a different category differed in their contributions
318 to exemplar representation from low face and word selective sites that did not demonstrate any
319 other category selectivity. The main question here is the extent to which contacts that
320 demonstrate category selectivity will contribute to exemplar representation for a different
321 category. Thus in addition to examining high and low category selective contacts, we further
322 decomposed the low face and word category selective populations into two sub groups: other
323 category selective (OCS) and not category selective (NCS). Category selectivity was established
324 with the same method as outlined above, and weights extracted in the same method as above.
325 The results of this analysis are depicted in Figure 5B and demonstrate that higher (but
326 comparable) proportions of NCS than OCS electrode contacts survive penalization and
327 contribute diagnostic information to exemplar classification using L1 regularization.
328

329 **Electrode diagnosticity as a continuous function of category selectivity**

330 To provide a non-binary depiction of the activity profiles of electrode contacts during face and
331 word individuation, we examined the relationships between the time course of electrode weight
332 magnitude and category selectivity. As a first comparison, peak *d'* value from the category
333 localizer was compared to the latency of peak weight value for each contact. This comparison
334 was conducted separately for face and word individuation. For clarity, only contacts with peak
335 weight values greater than 1 standard deviation above mean baseline (pre-stimulus) value were
336 included. Weight value in this instance, similar to the previous two paragraphs, refers to the
337 mean weight assigned to each contact for each of the pairwise exemplar comparisons at every

time point. High and low selectivity contacts (for faces and words, separately) were plotted together to demonstrate the general transition from early onset of non-zero weights of highly selective contacts to later onset of low selectivity contacts. The results of this analysis are depicted in Figure 4.

Statistical Analyses

For the category localizer with Naïve Bayes classification, row permutation tests on a subject level were used to establish a *d'* threshold for category selective contacts. For each subject within each permutation, the condition labels for each trial were randomly shuffled and the same classification procedure as above was used 1000 times for a randomly selected channel in each electrode montage. The peak *d'* value from each permutation was aggregated into a group-level distribution comprising the null distribution from each permutation for each subject. The *d'* value corresponding to $p < .01$ was estimated from this histogram and used as a selectivity threshold to determine high and low selectivity contact populations for each subject.

For face and word individuation as measured with elastic net regularized logistic regression, row permutation tests were used to establish a significance threshold for classification accuracy for each subject. For each permutation, a classifier model was optimized and test condition labels shuffled to test model predictions on randomized data. This procedure was repeated 1000 times to generate a null distribution. The true classification values and null distributions for each subject were combined into group-level distributions, and the mean true classification value and mean null distribution compared to one another. Classification accuracy was deemed significant at a level of $p < .05$ with FDR correction (Benjamini-Hochberg procedure for dependent tests), with a minimum temporal threshold of 3 contiguous significant

361 time points. Thus, although different subjects contributed different numbers of contacts to each
362 classification analysis, all subjects are weighted equally in the group mean depicted in Figure 2.
363 Onset sensitivity was determined by with 3 metrics examining the individual subject-
364 level statistics. For the first method, the same true classification values and null distributions
365 from above were compared on an individual level, and the first time point significant at a level of
366 $p < .05$ with FDR correction (Benjamini-Hochberg procedure for dependent tests) with a
367 minimum temporal threshold of 3 contiguous significant time points was used as the onset
368 marker for each subject. Vectors of onset markers compiled from all subjects were compared
369 between HFS / LFS, and HWS / LWS electrode populations with paired-sample t-tests. Because
370 this method is somewhat sensitive to the magnitude of the response (e.g. higher magnitude will
371 cross the statistical threshold sooner) two other methods for calculating onset that are more
372 robust to magnitude differences were used as well.
373 The second onset determination method was adapted from Schrouff et al. (2020): for each
374 subject, the time course of mean classification values for each classification problem (HFS, LFS,
375 HWS, and LWS) were normalized to peak classification value, and a sliding window with 50 ms
376 bins and 10 ms overlap was implemented. Classification average and standard deviation in the
377 baseline period of -100 to 0 ms was estimated, and the first period with 3 contiguous bins
378 surpassing the baseline threshold was marked as the signal onset for a given subject's
379 classification time course. Vectors of onset markers compiled from all subjects were compared
380 between HFS / LFS, and HWS / LWS electrode populations with paired-sample t-tests. Schrouff
381 et al (2020) show that this method for finding onset times is robust to differences in peak
382 magnitude across comparisons.

For the third onset determination method, onset sensitivity was measured as the first 3 contiguous time points where classification values for each subject were greater than 25% of the peak value. Vectors of onset markers compiled from all subjects were compared between HFS / LFS, and HWS / LWS electrode populations with paired-sample t-tests. While 25% of the peak value is not necessarily a strict measure of "onset," it is independent of peak magnitude and provides a metric of whether any differences in peak time are due to differences in slope or whether there is differences in onset (e.g. earlier peak times could be due to sharper rising slope or earlier onset).

RESULTS

Spatiotemporal dynamics of individuation

Significant face and word individuation was present in and out of HFS and HWS patches (Figure 2), as measured with elastic net regularized logistic regression. Using the first method of onset calculation, the onset of face individuation occurred 190 ms earlier in HFS patches relative to LFS patches ($t(13)= 3.05$, $p = 0.009$) and peaked 200 ms earlier ($t(13) = 2.73$, $p = 0.017$), with a higher peak in HFS than LFS patches ($t(13) = 2.68$, $p = 0.019$). Notably, the difference in the magnitude of the HFS and LFS response is independent of the difference in peak times, though onset times can be affected by magnitude differences. Using two other methods of onset calculation that are more robust to differences in magnitude (Schrouff et al., 2020), above chance face individuation occurred significantly earlier inside (160 ms, 210 ms) than outside (250 ms, 325 ms) HFS patches ($t(13) = 3.6$, $p = 0.003$; $t(13) = 3.03$, $p = 0.0096$).

Word individuation began 145 ms earlier in HWS patches relative to LWS patches ($t(4) = 3.1$, $p = 0.036$) and peaked 250 ms earlier ($t(4) = 3.61$, $p = 0.022$), with a higher peak in HWS

than LWS patches ($t(4) = 2.802$, $p = 0.048$). Using the two other methods of onset calculation that are more robust to differences in magnitude (Schrouff et al., 2020), above chance face individuation occurred earlier inside (150 ms, 190 ms) than outside (285 ms, 405 ms) HFS patches ($t(4) = 1.77$, $p = 0.15$; $t(4) = 4.31$, $p = 0.01$).

HFS and HWS patches maintained significant sensitivity to individual face and word representations respectively throughout visual processing (from 130-840 ms and 160-535 ms respectively, $p<0.05$ corrected for multiple comparisons), suggesting that these regions contribute to both early and late visual processing. LFS and LWS reached significance only later (from 320-800 ms and 285 - 605 ms respectively, $p < 0.05$ corrected for multiple comparisons), suggesting that these regions contribute to late visual processing. For both faces and words, the finding of earlier individuation in high selectivity regions relative to low selectivity regions was robust across a range of criteria for defining “high” and “low” selectivity (Figure 3). The robustness of the result demonstrates that illustrating that the differences in timing were not due to choosing an arbitrary threshold between high and low selectivity. Furthermore, see Figure 4 for a more continuous rather than binary examination of the relationship between selectivity and timing.

Electrodes were placed based on the clinical needs of the patients and not necessarily optimally placed for sensitivity to visual information, thus relative effect sizes are likely more relevant than absolute effect sizes. Peak effect sizes in LFS and LWS patches were relatively small, but nonetheless more than 1/3 that of the peak effect sizes in HFS and HWS patches. This suggests that activity in LFS and LWS patches contributed meaningfully to the overall representation of individual faces and words, albeit less than HFS and HWS patches.

429 **Relative contribution of high and low selectivity patches to individuation**

430 The previous results demonstrate that individuation emerges earlier in high selectivity than low
431 selectivity patches, but leaves the relative contribution of activity in high and low selectivity
432 patches to the overall individual-level representation unclear. Specifically, two important
433 questions are outstanding: 1) Is information in category low selectivity patches unique from
434 information in high selectivity patches? 2) If low selectivity patches contain unique information
435 that contributes to individuation, is this information present in patches selective to other
436 categories or patches that show no measured category selectivity, e.g. do word-selective contacts
437 contribute diagnostic information to face individuation?

438 Regarding the first question, it is possible that the later, above-chance individuation
439 accuracy in low selectivity patches is solely due to activity that is highly correlated with activity
440 from high category-selective patches. If so, this would suggest the information in the low
441 selectivity patches does not contribute additional information to face or word individuation at the
442 whole-brain level. To test this hypothesis, sparse classification using L1-regularization and
443 identical parameters to earlier elastic net procedure except regularization parameter (λ) was
444 performed over all ventral temporal contacts to identify the electrode contacts that provided
445 information for face or word individuation. If activity between any set of contacts is highly
446 correlated, L1-regularization should force all contacts in that set to have zero weight, except the
447 one with the largest amount of discriminating information. Thus, if low selectivity contacts
448 convey redundant information to that in high selectivity contacts, low selectivity contacts should
449 be penalized and removed from the model, given that individuation was weaker for low
450 selectivity compared to high selectivity contacts (Figure 2). However, if low selectivity contacts
451 do contain unique diagnostic information that contributes to individuation, a certain proportion

452 of low selectivity contacts should be assigned non-zero weights. Note that choice of
453 regularization method (elastic net vs. L1) does not alter the pattern of reported results above. To
454 address the second question, the above analysis was extended by decomposing the low
455 selectivity contacts into "other category-selective" and "not category-selective" populations. This
456 was done by identifying the LFS and LWS contacts that showed high selectivity for any of the
457 other 5 categories in the localizer and those that did not.

458 For both face and word individuation tasks, the analysis showed that proportions of both
459 high selectivity and low selectivity electrode populations contribute diagnostic information
460 (Figure 5A), though high selectivity patches may contribute more than low selectivity ones.
461 Second, decomposing the low selectivity contacts showed that in the face individuation task,
462 regions highly selective for other categories contribute diagnostic information to overall
463 individuation as well as those that demonstrate low selectivity for all categories (Figure 5B).
464 These findings demonstrate that in the later time period of significant individuation in low
465 selectivity contacts, meaningful information that contributed to above chance individuation is
466 present outside of category-selective areas, distributed even to areas that demonstrate selectivity
467 for a different visual object category.

468

469 **Discussion**

470 The presence of individual-level information in and out of highly category-selective areas at
471 different latencies suggests a "dynamic domain specificity" model of visual processing.
472 Specifically, information from a given visual category is first processed primarily in strongly
473 category-selective patches followed by widespread processing that includes both patches that are
474 strongly and weakly selective for that stimulus category (Shehzad & McCarthy, 2018) The

475 cascade of neural activity during visual perception is characterized by an early, potentially
476 obligatory, stage of processing in strongly category-selective patches that may guide and gate
477 information for further processing. Previous studies suggest that this early stage represents a
478 coarse pass of processing only allowing for differentiation of relatively distinct images (Hirshorn
479 et al., 2016; Ghuman et al., 2014; Hegdé, 2008). Approximately 150-200 ms later, information
480 then flows to visual processing patches outside of strongly category-selective patches as well,
481 including into patches that are selective for other visual categories, either through lateral and
482 recurrent connectivity or through top-down feedback. Low selectivity patches contribute unique
483 information to the overall individual-level representation (Figure 5) in the later time period
484 (Figure 2), albeit somewhat less information than highly selective patches. Thus, low selectivity
485 patches may help support later visual processing (Hirshorn et al, 2016; Ghuman et al., 2014; Li
486 et al., 2019) that could contribute to determining subtle distinctions between individual category
487 members or assist with later processes such as viewpoint or position generalization (Freiwald et
488 al., 2010). The contribution of low selectivity patches may be particularly important for
489 perception under challenging conditions, such as occlusion or otherwise degraded conditions. A
490 recent iEEG showed that there was little difference in the dynamics of high and low selectivity
491 regions at the category level and that the category-level information in low selectivity regions
492 was redundant with the information in high selectivity regions (Schrouff et al., 2020). The
493 current results focusing on individual-level representation suggest that highly category-selective
494 patches contribute to the neural representation in both early and later processing stages, and low
495 selectivity patches provide non-redundant information that support later processing stages.
496 Regarding the spread of high selectivity contacts throughout VTC, our findings are comparable
497 to the results of other iEEG studies mapping category selectivity which reveal selective contacts

located on, lateral, and medial to the fusiform gyrus (Kadipasaoglu et al., 2016; Allison et al., 1999) and are not spatially identical to group-averaged categorical topographies as measured with fMRI.

The proposed dynamic domain specificity hypothesis may reconcile apparent contradictions between findings that have been used to support domain-specific and distributed feature map models of visual perception. The profound and frank disturbances to the perception of stimuli from particular categories seen in the presence of lesions or disruptions to highly category-selective patches (Puce et al., 1999; Parvizi et al., 2012; Afraz et al., 2006; Farah et al., 1995; Schalk et al., 2017) may emerge due to the perturbation of early and potentially obligatory activity of these areas during visual processing. The perceptual relevance of later activity in low selectivity patches is supported by the current evidence that these patches contribute unique information to face and word individuation (Figure 5). The time of peak individuation in low selectivity patches occurs when significant individuation is still present in high selectivity patches and is near the time when key higher-level visual processes such as viewpoint generalization (Freiwald et al., 2010) and semantic processing (Clarke et al., 2015) occur. Additionally, single units in the medial temporal lobes show selectivity for individual faces in a similar later time period and it has been suggested that this time period is critical for linking perception and memory (Quian Quiroga, 2012; Mormann et al., 2008). Furthermore, this time window is substantially earlier than behavioral reaction times for comparable individual-level face and word recognition tasks (Haxby et al., 1999; Seidenberg & McClelland, 1989). The later information processing in low selectivity patches would also help explain why category discriminant information is sometimes seen outside of category-selective patches in low temporal resolution measures such as fMRI (Haxby et al., 2001; Ghuman & Martin, 2019). As

521 such, low selectivity patches may play a role in some aspects of individuation, even if that role is
522 later and more supportive than the central role of strongly selective patches. Causal evidence is
523 required to test whether activity outside of strongly selective patches contributes to perception;
524 for example, an alternative explanation of later discrimination in these regions is that it could
525 reflect a backpropagating learning signal (Rumelhart et al., 1986) rather than perceptual
526 processing.

527 While the results here are consistent with the primarily low temporal resolution data that
528 have been used to support both domain specific and distributed feature map models of VTC
529 organization, they also help address theoretical aspects of the debate between the models.
530 Specifically, in distributed feature map models the difference between strongly and weakly
531 selective parts of VTC is a difference in the degree to which each contributes to perception of
532 stimuli from a particular category, but these contributions should happen at the same processing
533 stage. These models would predict that strongly and weakly selective regions should each have
534 similar timecourses of processing, varying mostly in how much each contributes to the
535 representation for a particular stimulus class. The result that individual-level representations in
536 highly selective regions onset and peak 145 - 250 ms earlier than in weakly selective regions
537 presents a challenge to current instantiations of distributed feature map models. These
538 differences survive across a range of criteria for selectivity (Figure 3), suggesting there is a
539 qualitative, not graded, difference in the role that highly selective regions play for processing
540 stimuli that those regions are selective for relative to the rest of VTC. Thus, continuous feature
541 map models would need to be modified to accommodate relationships between selectivity and
542 latency of information processing.

In the strongest versions of domain specificity models, there is no role for regions of VTC weakly selective for a particular category of image in perceptual processing for that stimulus type. However, the results here suggest that these weakly selective regions do contribute to later visual processing. The dynamic domain specificity hypothesis outlined above is an attempt to modify traditional models of domain specificity by positing a supportive role for weakly selective regions that may support later processes and perhaps aid in perception under challenging perceptual conditions.

The dynamic pattern of results was seen for both faces, with circuitry that putatively arises from evolutionary and genetic origins, and words, where reading skill must be acquired fully through experience, suggesting dynamic domain specificity may be a general principle of cortical organization. Taken together, these results may reconcile the tension between domain-specific versus distributed feature map models of visual object processing by providing evidence that domain-specific and distributed processing emerge dynamically at different times during the course of visual perception.

Author Contributions: A.S.G designed the experiment. R.M.R. conducted surgical implantations. B.B.B., M.J.B., R.M.R., and A.S.G. collected experimental data. B.B.B. and M.J.B. analyzed the data. B.B.B. and A.S.G. wrote the manuscript.

Data and materials availability: All data and code is available upon reasonable request to A.S.G (ghumana@upmc.edu).

580 **References**

581 Afraz SR, Kiani R, Esteky H (2006) Miscrostimulation of inferotemporal cortex influences face
582 categorization. Nature 442:692-695.

583 Allison T, Puce A, Spencer DD, McCarthy G (1999) Electrophysiological studies of human face
584 perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. Cereb
585 Cortex 9:415-430.

586 Balota, DA, Yap MJ, Hutchison KA, Cortese MJ, Kessler B. Loftis B, Neely JH, Nelson DL,
587 Simpson GB, Treiman R (2007) The English Lexicon Project. Beh. Res. Methods 39:445–459.

588 Bell, A. H., Malecek, N. J., Morin, E. L., Hadj-Bouziane, F., Tootell, B. H. & Ungerleider, L. G.
589 (2011) Relationship between fMRI-identified regions and neuronal category-selectivity. J.
590 Neurosci. 3:12229 – 12240.

591 Brainard DH (1997) The Psychophysics Toolbox. Spat. Vis. 10:433-436.

592 Clarke A, Devereux BJ, Randall B, Tyler LK (2015) Predicting the time course of individual
593 objects with MEG. Cerebr. Cortex 10:3602 – 3612.

594 Cohen J (1988) Statistical Power Analysis for the Behavioral Sciences, ed. 2. UK: Routledge.

595 Cukur T, Huth AG, Nishimoto S, Gallant JL (2013) Functional subdomains within human FFA.
596 J. Neurosci. 33:16748 – 16766.

597 Dale AM, Fischl B, Sereno MI (1999) Cortical surface-based analysis. I. Segmentation and
598 surface reconstruction. NeuroImage 9:179 – 194.

599 Damasio AR, Damasio H, Van Hoesen GW (1982) Prosopagnosia: anatomic basis and
600 behavioral mechanisms. Neurol. 32:331 – 41.

601 Farah MJ, Levinson KL, Klein KL (1995) Face perception and within-category discrimination in
602 prosopagnosia. Neuropsychologia 33:661-674.

603 Freiwald WA, Tsao DY (2010) Functional compartmentalization and viewpoint generalization
604 within the macaque face-processing system. Science 330:845 – 851.
605 Ghuman AS, Brunet NM, Li Y, Konecky RO, Pyles JA, Walls SA, Destefino V, Wang W,
606 Richardson RM (2014) Dynamic encoding of face information in the human fusiform gyrus. Nat.
607 Comm. 5:5672.
608 Ghuman AS, Martin A (2019) Dynamic neural representations: An inferential challenge for
609 fMRI. Trends Cogn. Sci. 23:534 – 536.
610 Hanson SJ, Schmidt A (2011) High-resolution imaging of the fusiform face area (FFA) using
611 multivariate non-linear classifiers shows diagnosticity for non-face categories. NeuroImage
612 54:1715 – 1734.
613 Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P (2001) Distributed and
614 overlapping representations of faces and objects in ventral temporal cortex. Science 293:2425-
615 2430.
616 Haxby JV, Ungerleider LG, Clark V, Schouten J, Hoffman E, Martin A (1999) The effect of face
617 inversion on activity in human neural systems for face and object perception. Neuron 22:189 –
618 99.
619 Hegdé J (2008) Time course of visual perception: coarse-to-fine processing and beyond. *Prog.*
620 Neurobiol. 84:405 - 439.
621 Hermes D, Miller KJ, Noordmans HJ, Vansteensel MJ, Ramsey NF (2010) Automated
622 electrocorticographic electrode localization on individually rendered brain surfaces. J. Neurosci.
623 Methods 185:293-298.

624 Hirshorn EA, Li Y, Ward MJ, Richardson RM, Fiez JA, Ghuman AS (2016) Decoding and
625 disrupting left midfusiform gyrus activity during word reading. Proc. Natl. Acad. Sci. USA
626 113:8162 – 8162.
627 Kadipasaoglu CM, Conner CR, Whaley ML, Baboyan VG Tandon N (2016) Category-selectivity
628 in human visual cortex follows cortical topology: A grouped icEEG study. Plos One.
629 Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: A module in human
630 extrastriate cortex specialized for face perception. J. Neurosci. 17:4302 – 4311.
631 Kiani R, Esteky H, Mirpour K, Tanaka K (2007) Object category structure in response patterns
632 of neuronal population in monkey inferior temporal cortex. J. Neurophysiol. 97:4296 – 4309.
633 Leszczynski M, Barczak A, Kajikawa Y, Ulbert I, Falchier A, Tal I, Haegens S, Melloni L,
634 Knight R, Schroeder C (2019) Dissociation of broadband high-frequency activity and neuronal
635 firing in the neocortex. bioRxiv. https://doi.org/10.1101/531368
636 Li Y, Richardson RM, Ghuman AS (2019) Posterior fusiform and midfusiform contribute to
637 distinct stages of facial expression processing. Cereb. Cortex 186.
638 Livingstone MS, Arcaro MJ, Schade PF (2019) Cortex is cortex: Ubiquitous principles drive
639 face-domain development. Trends Cogn. Sci. 23:3 – 4.
640 Martin A (2007) The representation of object concepts in the brain. Annu. Rev. Psycho. 58:25-
641 45.
642 Miller KJ, Schalk G, Hermes D, Ojemann JG, Rao RP (2016) Spontaneous decoding of the
643 timing and content of human object perception from cortical surface recordings reveals
644 complementary information in the event-related potential and broadband spectral change. PLOS
645 Comput. Biol. 12.

Mormann F, Kornblith S, Quian Quiroga R, Kraskov A, Cerf M, Fried I, Koch C (2008) Latency and selectivity of single neurons indicate hierarchical processing in the human medial temporal lobe. J. Neurosci. 28**:**8865 – 8872.

Mur M, Ruff DA, Bodurka J, De Weerd P, Bandettini PA, Kriegeskorte N (2012) Categorical, yet graded – single-image activation profiles of human category-selective cortical regions. J. Neurosci, 32:8649 – 8662.

Nestor A, Behrmann M, Plaut DC (2011) Unraveling the distributed neural code of facial identity through spatiotemporal pattern analysis. Proc. Natl. Acad. Sci. USA 108:9998 – 10003.

Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. 156869.

Op de Beeck HP, Haushofer J, Kanwisher N (2008) Interpreting fMRI data: maps, modules and dimensions. Nat. Rev. Neurosci. 9**:**123-135.

Parvizi J, Jacques C, Foster BL, Withoft N, Rangarajan V, Weiner KS,Grill-Spector K (2012) Electrical stimulation of human fusiform face-selective regions distorts face perception. J. Neurosci. 32:14915 – 14920.

Perrett DI, Smith PA, Potter DD, Mistlin AJ, Head AS, Milner AD, Jeeves AD (1984) Neurones responsive to faces in the temporal cortex: Studies of functional organization, sensitivity to identity and relation to perception. Hum. Neurobiol. 3:197 – 208.

Pitcher D, Walsh V, Yovel G, Duchaine B (2007) TMS evidence for the involvement of the right occipital face area in early face processing. Curr. Biol. 17:1568 – 1573.

Puce A, Allison T, Asgari M, Gore JC, McCarthy G (1996) Differential sensitivity of human visual cortex to faces, letterstrings, and textures: A functional magnetic resonance imaging study. J. Neurosci. 16, 5205 - 5215.

Puce A, Allison T, McCarthy G (1999) Electrophysiological studies of human face perception. III: Effects of top-down processing on face-specific potentials. Cereb. Cortex 9:445 – 458.

Quian Quiroga R (2012) Concept cells: the building blocks of declarative memory functions. Nat. Rev. Neurosci. 13:587 – 597.

Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. Nature 323:533-536.

Schalk G, Kapeller C, Guger C, Ogawa H, Hiroshima S, Lafer-Sousa R, Saygin ZM, Kamada K, Kanwisher N (2017) Facephenes and rainbows: Causal evidence for functional and anatomical specificity of face and color processing in the human brain. Proc. Natl. Acad. Sciences 114: 12285-12290.

Seidenberg MS, McClelland JL (1989) A distributed, developmental model of word recognition and naming. Psychol. Rev. 96:523 - 568.

Shehzad Z, McCarthy G (2018) Category representations in the brain are both discretely localized and widely distributed. J. Neurophysiol. 119:2256 – 2264.

Spiridon M, Kanwisher N (2002) How distributed is visual category information in human occipito-temporal cortex? An fMRI study. Neuron 35:1157 – 1165.

Tadel F, Baillet S, Mosher JC, Pantazis D, Leahy RM (2011) Brainstorm: a user-friendly application for MEG/EEG analysis. Comput. Intell. Neurosci. 8.

Tsao DY, Freiwald WA, Tootell RB, Livingstone MS (2006) A cortical region consisting entirely of face-selective cells. Science 311:670 - 674.

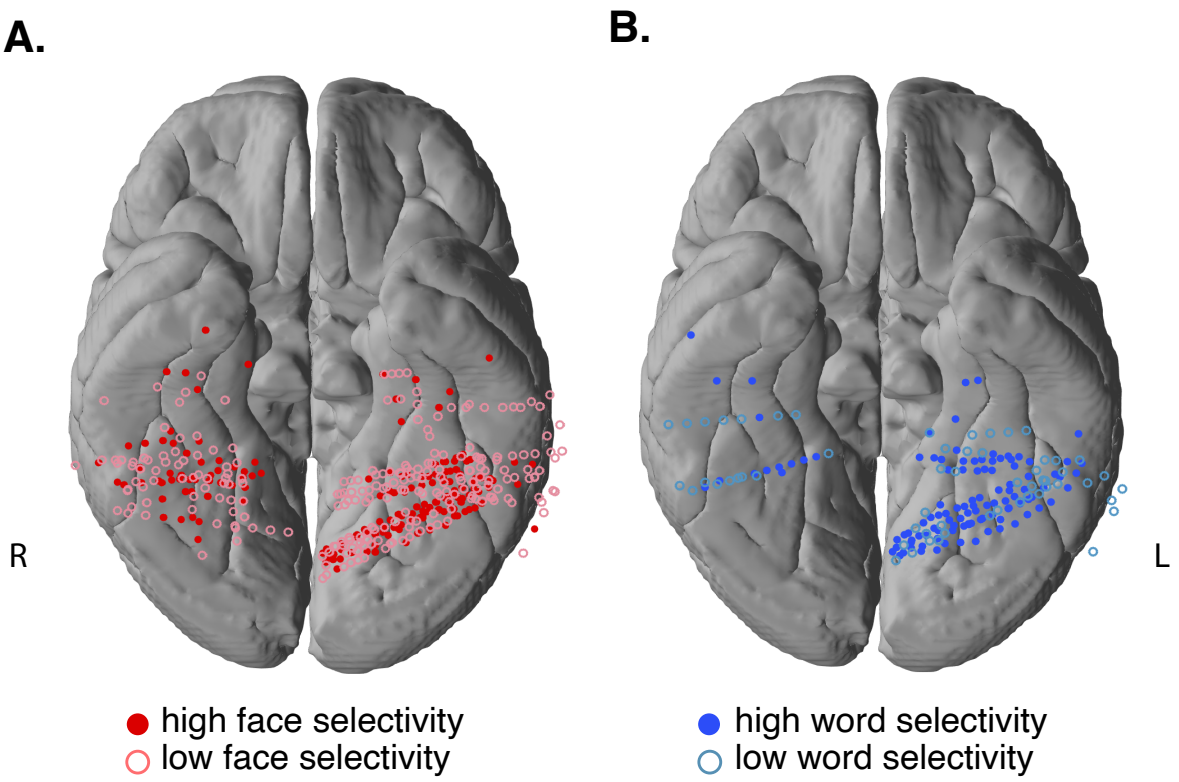

Figure 1. Ventral temporal electrode contacts (489 total). Electrode contacts depicted here include both subdural electrode strips on the surface of the cortex and stereotaxically implanted depth strips, projected to the nearest surface vertex. A) 426 total contacts from 14 subjects were divided into 171 high face selectivity and 255 low face selectivity contacts. B) 174 total contacts from 5 subjects were divided into 113 high word selectivity and 61 low word selectivity contacts.

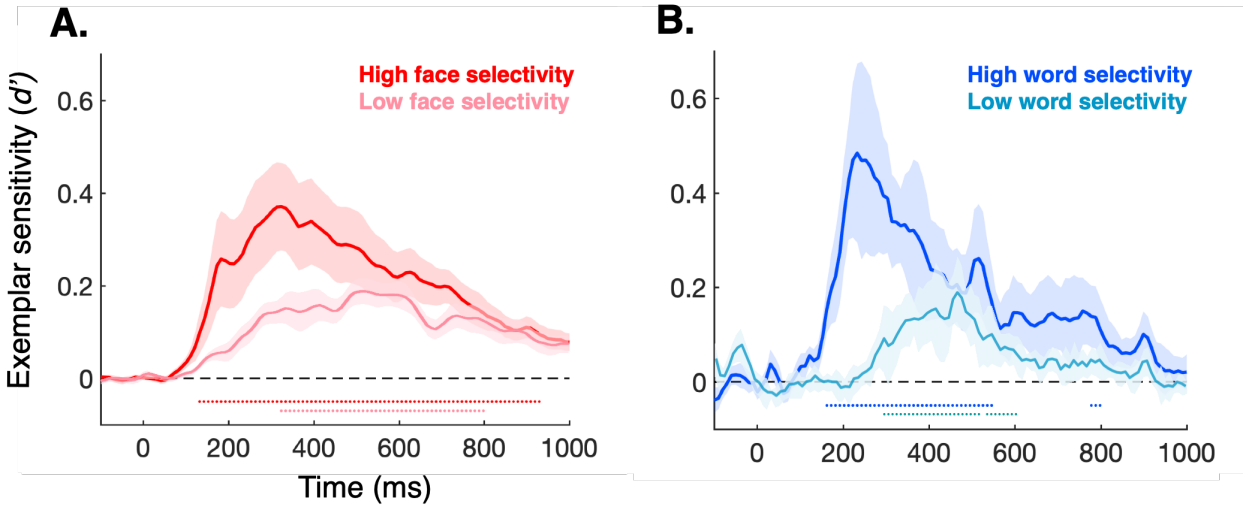

Figure 2. Time course of the sensitivity index (*d'*) for individuation of faces and words. A) HFS contacts (dark red) demonstrated significant face individuation from 130 ms to 840 ms after stimulus onset, with peak *d'* = 0.37 at 320 ms ($p < .05$, FDR corrected for multiple dependent temporal comparisons). LFS contacts (light red) demonstrated significant face individuation from 320 ms to 800 ms after stimulus onset, with peak *d'* = 0.17 at 520 ms ($p < .05$, FDR corrected). HFS individuation onset emerged significantly earlier than LFS individuation (190 ms average difference between HFS and LFS onset, $p = .009$, $t(13) = 3.05$). Individually, 11 of the 14 patients demonstrated an earlier onset of significant individuation in HFS than LFS contacts. B) HWS contacts (dark blue) demonstrated significant word individuation from 160 ms to 535 ms after stimulus onset, with peak *d'* = 0.48 at 235 ms ($p < .05$, FDR corrected for dependent tests). LWS contacts (light blue) demonstrated significant word individuation from 285 ms to 605 ms after stimulus onset, with peak *d'* = 0.18 at 470 ms ($p < .05$, FDR corrected). Word individuation emerged significantly earlier in HWS compared to LWS patches (145 ms average difference between HWS and LWS onset, $p = .036$, $t(4) = 3.1$). Individually, all 5 patients demonstrated an earlier onset of word individuation in HWS compared to LWS contacts. Shaded bars illustrate standard error of the mean across subjects at each time point

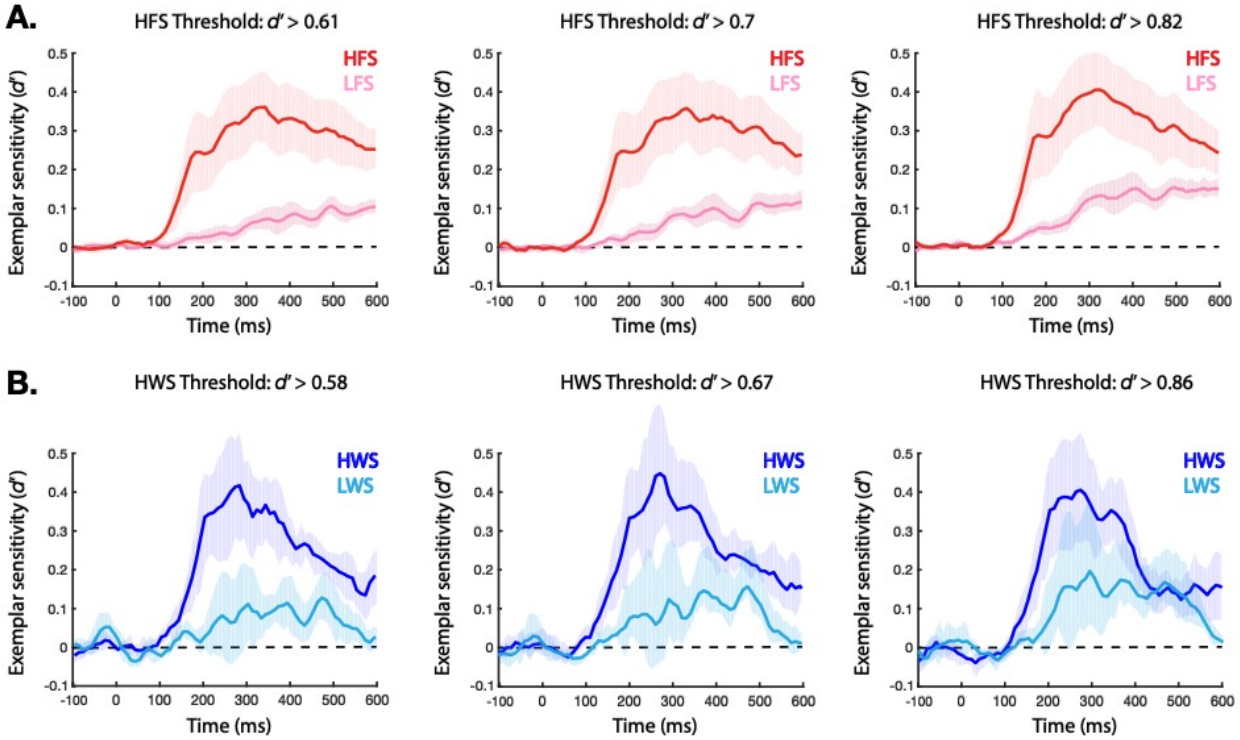

Figure 3. Time course of the sensitivity index (*d'*) for individuation of faces and words at graded thresholds. Individual face and word classification (as in Figure 3) was repeated with multiple "high" and low" selectivity thresholds. These thresholds were defined by separating all contacts into partitions corresponding to the one third, one half, and two-thirds levels of d' values as measured with the category localizer. A) Time course of face individuation at thresholds of d' = 0.61 (LFS = bottom 1/3 of contacts, HFS = top 2/3 of contacts), d' = 0.7 (LFS = bottom ½ of contacts, HFS = top ½ of contacts), and d' = 0.82 (LFS = bottom 2/3 of contacts, HFS = top 1/3 of contacts). B) Time course of word individuation at thresholds of d' = 0.58 (LWS = bottom 1/3 of contacts, HWS = top 2/3 of contacts), d' = 0.67 (LWS = bottom ½ of contacts, HWS = top ½ of contacts), and d' = 0.86 (LWS = bottom 2/3 of contacts, HWS = top 1/3 of contacts).

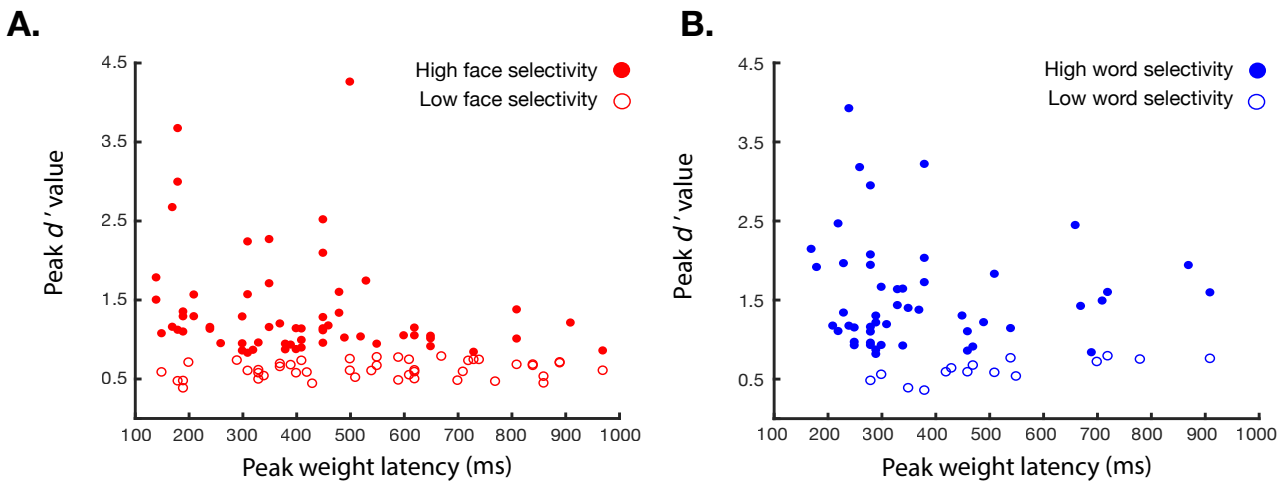

Figure 4. Relationship between selectivity (*d*') and latency of exemplar information. The selectivity metric was extracted from the category localizer used to define HFS/HWS and LFS/LWS populations. Exemplar information as measured with peak weight value refers to the mean weight assigned to each contact for each of the pairwise exemplar comparisons at every time point for the L1-regularized classifier used for Figure 2. A) Scatter plot depicting latency of peak weight value for HFS and LFS contacts by peak category selectivity value. For this depiction, only contacts with peak values > 1 standard deviation above mean group baseline weight value were included because in electrodes with little diagnostic signals peak times are unreliable and mostly random. 20 out of 62 HFS electrodes have peaks and 6 out of 46 LFS electrodes peak prior to 300 ms. R^2 between *d'* and peak weight latency = 0.11 ($p = 0.006$). B) Scatter plot depicting latency of peak weight value for HWS and LWS contacts by peak category selectivity value (same thresholding as A). 26 out of 50 HWS electrodes have peaks and 2 out of 15 LWS electrodes peak prior to 300 ms. R^2 between *d'* and peak weight latency = 0.04 ($p = 0.11$). This set of results shows that both for word and face processing, more selective electrode contacts have earlier peak weight latencies than less selective electrode contacts. These data speak to a relationship between category selectivity and latency of face or word individuation.

However, the existence of inter-subject variability and weak correlation means we cannot rule out a graded or binary relationship among electrode contacts based on degree of category selectivity as some correlation would be expected in either case.

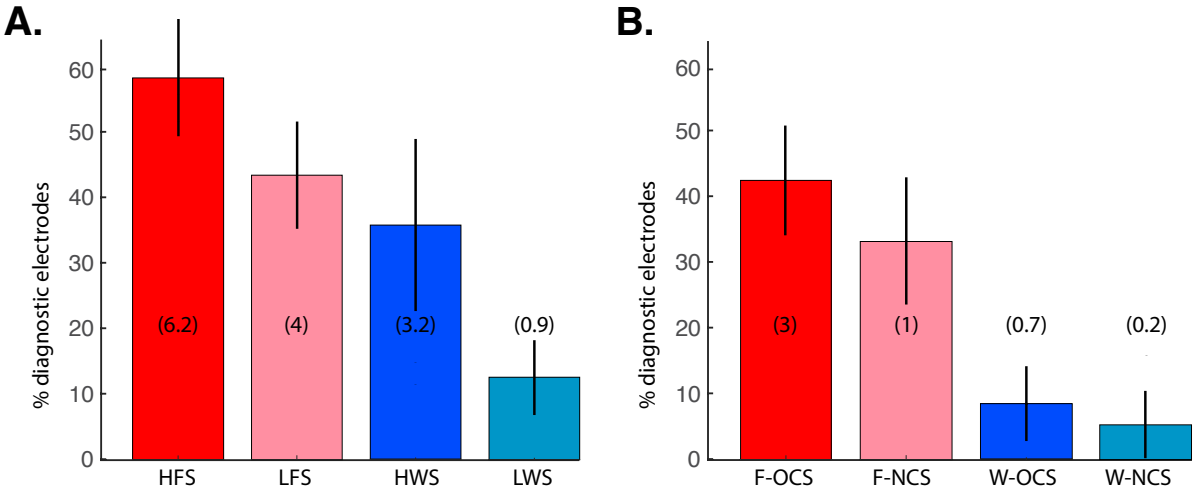

Figure 5. A) Percent diagnostic electrode contacts in high category-selectivity and low selectivity populations measured with L1-regularized logistic regression. 58.6% of HFS contacts from 14 patients (average of 6.2 contacts per subject, SEM = 9%) were assigned non-zero weights. 42.4% of LFS contacts (average of 4 contacts, SEM = 8.9%) were assigned non-zero weights on average. 35.2% of HWS contacts from 5 patients (average of 3.2 contacts, SEM = 18%) on average were assigned non-zero weights. 11.4% of LWS contacts (average of 0.9 contacts, SEM = 6.2%) were assigned non-zero weights. B) Within low face selectivity electrode contacts that demonstrated selectivity for a different object category (F-OCS, other category-selective), 42.5% of the contacts (average of 3 contacts per montage, SEM = 8.7% of contacts) were assigned non-zero weights. Within low face selectivity contacts that demonstrated no selectivity for any object categories (F-NCS, not category-selective), 33% of the contacts (average of 1 contact, SEM = 9.6%) on average were assigned non-zero weights. Within low word selectivity electrode contacts that demonstrated selectivity for a different object category (W-OCS), 9.4% of the contacts (average of 0.7 contacts, SEM = 8.1%) were assigned non-zero weights. Within low word selectivity contacts that demonstrated low selectivity for any other

object categories (W-NCS), 5% of the contacts (average of 0.2 contacts, SEM 5%) on average were assigned non-zero weights.